PCT/AU2004/000856

AUSTRALIA

REC'D 13 JUL 2004
WIPO PCT

Patent Office
Canberra

I, JULIE BILLINGSLEY, TEAM LEADER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2003903317 for a patent by PROTEOME SYSTEMS INTELLECTUAL PROPERTY PTY LTD as filed on 27 June 2003.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

WITNESS my hand this
Eighth day of July 2004

J. Billingsley

JULIE BILLINGSLEY
TEAM LEADER EXAMINATION
SUPPORT AND SALES

AUSTRALIA

Patents Act 1990

Proteome Systems Intellectual Property Pty Ltd

PROVISIONAL SPECIFICATION

Invention Title:

Method of isolating a protein

The invention is described in the following statement:

Field of the invention

The present invention relates to a method of identifying an immunogenic protein isolated in a biological sample obtained from a subject. In particular, the immunogenic protein is isolated and/or identified in an immunoglobulin fraction obtained from a subject. This method is of particular use in obtaining novel diagnostic/prognostic/therapeutic target molecules for the treatment or diagnosis of diseases, such as, for example, those caused by infectious organisms or autoimmune disease.

Background of the invention

General Information

As used herein the term "derived from" shall be taken to indicate that a specified integer may be obtained from a particular source albeit not necessarily directly from that source.

Unless the context requires otherwise or specifically stated to the contrary, integers, steps, or elements of the invention recited herein as singular integers, steps or elements clearly encompass both singular and plural forms of the recited integers, steps or elements.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements or integers but not the exclusion of any other step or element or integer or group of elements or integers.

Unless specifically stated otherwise, each feature described herein with regard to a specific aspect or embodiment of the invention, shall be taken to apply *mutatis mutandis* to each and every other aspect or embodiment of the invention.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations or any two or more of said steps or features.

The present invention is not to be limited in scope by the specific examples described herein. Functionally-equivalent products, compositions and methods are clearly within the scope of the invention, as described herein.

All the references cited in this application are specifically incorporated by reference herein.

The present invention is performed without undue experimentation using, unless otherwise indicated, conventional techniques of molecular biology, microbiology, virology, recombinant DNA technology, peptide synthesis in solution, solid phase peptide synthesis, and immunology. Such procedures are described, for example, in the following texts that are incorporated by reference:

Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, New York, Second Edition (1989), whole of Vols I, II, and III;

DNA Cloning: A Practical Approach, Vols. I and II (D. N. Glover, ed., 1985), IRL Press, Oxford, whole of text;

Oligonucleotide Synthesis: A Practical Approach (M. J. Gait, ed., 1984) IRL Press, Oxford, whole of text, and particularly the papers therein by Gait, pp1-22; Atkinson *et al.*, pp35-81; Sproat *et al.*, pp 83-115; and Wu *et al.*, pp 135-151;

Nucleic Acid Hybridization: A Practical Approach (B. D. Hames & S. J. Higgins, eds., 1985) IRL Press, Oxford, whole of text;

Immobilized Cells and Enzymes: A Practical Approach (1986) IRL Press, Oxford, whole of text;

Perbal, B., A Practical Guide to Molecular Cloning (1984);

Methods In Enzymology (S. Colowick and N. Kaplan, eds., Academic Press, Inc.), whole of series;

J.F. Ramalho Ortigão, "The Chemistry of Peptide Synthesis" *In:* Knowledge database of Access to Virtual Laboratory website (Interactiva, Germany);

Sakakibara, D., Teichman, J., Lien, E. Land Fenichel, R.L. (1976). *Biochem. Biophys. Res. Commun.* 73 336-342

Merrifield, R.B. (1963). *J. Am. Chem. Soc.* 85, 2149-2154.

Barany, G. and Merrifield, R.B. (1979) in *The Peptides* (Gross, E. and Meienhofer, J. eds.), vol. 2, pp. 1-284, Academic Press, New York.

Wünsch, E., ed. (1974) *Synthese von Peptiden in Houben-Weyls Metoden der Organischen Chemie* (Müler, E., ed.), vol. 15, 4th edn., Parts 1 and 2, Thieme, Stuttgart.

Bodanszky, M. (1984) *Principles of Peptide Synthesis*, Springer-Verlag, Heidelberg.

Bodanszky, M. & Bodanszky, A. (1984) *The Practice of Peptide Synthesis, Springer-Verlag*, Heidelberg.

Bodanszky, M. (1985) *Int. J. Peptide Protein Res.* 25, 449-474.

Handbook of Experimental Immunology, Vols. I-IV (D. M. Weir and C. C. Blackwell, eds., 1986, Blackwell Scientific Publications).

Description of the related art

As a response to the increasing demand for new diagnostic targets, lead compounds and new target identification and validation reagents, the pharmaceutical industry has increased its screening for new markers or compounds specific to pathogenic organisms or disease states, such as, for example, in the diagnosis/prognosis and/or treatment of infection and autoimmune disease.

As many pathogenic organisms express unique proteins or isoforms of proteins, much research has focussed on identifying and using these proteins for use in the development of novel diagnostic/prognostic and/or therapeutic strategies. However, not all proteins expressed by a pathogenic organism represent suitable targets for use in a method of diagnosis or in a therapeutic strategy. Accordingly, a large amount of the research in identifying new diagnostic and/or therapeutic strategies is directed toward the identification of suitable target molecules.

Perhaps the simplest approach in identifying a diagnostic/prognostic target of interest is to determine a protein derived from a pathogen that is associated with a disease or condition, or alternatively, to determine a host cell protein having an altered expression pattern as a consequence of the disease or conditions. The determined protein is then used to generate or identify an antibody, ligand or small molecule that is able to specifically bind to said protein or a region thereof, to determine whether or not the protein is sufficiently immunogenic to facilitate its use in the preparation of immuno-diagnostic reagents. However, such methods experience a high rate of failure because, many of the proteins that are tested are not immunogenic or at least not to the degree required to elicit an immune response in a host, for the production of immuno-diagnostic reagents and kits or vaccines.

Furthermore, a target of an antibody, ligand or small molecule may be relatively inaccessible in the native environment, ie in a complex with other proteins or within a cell, thereby hindering its detection by immunoassay.

Accordingly, the high failure rate of such a method means that this approach is both laborious and expensive, as often several potential targets must be tested before a putative target is identified.

With the completion of the sequencing of the genome of several pathogenic organisms researchers have commenced using this information to attempt to predict the function of proteins that are expressed by these pathogenic organisms. Using both function and sequence information researchers attempt to predict the location and accessibility of proteins expressed by the pathogenic organism, and thus the likelihood that a protein represents a diagnostic or therapeutic target for the treatment of an infection. As reported by Masignani *et al, Expert Opin. Biol. Ther. 2(8),* 895-905, 2002, this process can lead to the rapid prediction of putative diagnostic, therapeutic and/or vaccine targets, leading to an acceleration of the development of new therapeutic/diagnostic opportunities.

Methods that depend upon analysis of the genome sequence of an organism require that the genome, or at least a significant proportion of the genome of the organism of interest be sequenced. Accordingly, this method is ineffective at predicting diagnostic/therapeutic targets in organisms that have genomes that are yet to be sequenced, especially in those organisms that have only been recently identified. Furthermore, such an approach is of limited use in the prediction of potential therapeutic/diagnostic targets in pathogens, such as, for example, some retroviruses, which maintain a high mutation rate, thus regularly changing their genomic sequence.

Furthermore, these methods require the skilled artisan to determine which, if any, of the predicted target proteins are actually expressed by the pathogenic organism in vivo and during infection of a host cell.

As can be perceived from the preceding discussion, there remains a need for a method of isolation and/or identification of diagnostic target proteins.

Summary of the Invention

In work leading up to the present invention the inventors sought to identify novel diagnostic, vaccine and drug target proteins for infection by *Mycobacterium tuberculosis.* The inventors found that they could recover immunogenic target proteins or peptide fragments from the Immunoglobulin (Ig) fraction of patient serum samples, which proteins or peptide fragments were sufficiently non-degraded to permit determination of their amino acid sequences. This was surprising since proteins are known to be rapidly degraded during infection. Accordingly, the inventors proceeded against conventional wisdom in the art, by identifying several proteins/fragments from the Ig fraction of serum obtained from patients suffering from tuberculosis. The approach taken by the inventors is of general use in the identification of any immunogenic protein for use in diagnostic applications to identify pathogenic organism or infectious state in a subject. Additionally, such a protein is also of use in developing therapeutic or prophylactic strategies for the treatment of an infection by a pathogenic organism from which said protein is isolated.

Accordingly one aspect of the present invention provides a method of identifying an immunogenic protein or fragment thereof in a subject capable of eliciting an immune response to the immunogenic protein or fragment thereof, said method comprising:

(i) obtaining an immunoglobulin fraction from the subject; and
(ii) identifying a protein or fragment thereof, in said immunoglobulin fraction, other than an immunoglobulin protein or fragment thereof.

Preferably, the method of the invention comprises obtaining a sample of blood, serum or plasma from the subject and/or producing an immunoglobulin fraction thereof.

For example, the present inventors have isolated a 49.7 kDa protein derived from *M. tuberculosis* from an immunoglobulin fraction isolated from a serum sample isolated from a subject suffering from TB. Based upon sequence analysis this protein appears to be a glutamine synthetase. Accordingly, the inventors have used method of the present invention to definitively demonstrate that this previously predicted protein is expressed by *M. tuberulosis,* and furthermore is detectable in a patient sample.

In one embodiment, an "immunoglobulin fraction" shall be taken to include any purified or partially purified fraction of a biological sample that comprises an immunoglobulin-producing cell, or immunoglobulin, such as, for example, IgM, IgG,

IgA, IgE or IgD in addition to any cell or protein that is isolated with an immunoglobulin, such as, for example, a component of the complement pathway or an antigen. Accordingly, an antibody-containing fraction clearly includes an IgM or IgG, IgA, IgE or IgD fraction or peptide fragment thereof isolated from plasma.

Preferably, an immunoglobulin fraction is purified without separating or purifying a free immunoglobulin fraction from an immunoglobulin that is associated with the immunogenic protein. As used herein the term "free immunoglobulin" shall be taken to mean an immunoglobulin that is not complexed with another molecule, and in particular an antigen, immunogenic protein or fragment thereof

In one embodiment, the method of the present invention identifies a native or uncleaved form of an immunogenic protein or fragment thereof that is isolated with an immunoglobulin fraction from a biological sample derived from a patient.

In another embodiment, the method of the present invention identifies a cleaved form, a fragment, a peptide or an epitope of an immunogenic protein that is obtained with an immunoglobulin fraction from a biological sample derived from a patient.

In one embodiment, the method of the present invention further comprises the step of dissociating an immunogenic protein or fragment thereof that is obtained with an immunoglobulin fraction from an immunoglobulin.

As a further step a sample comprising an immunoglobulin and an immunogenic protein or fragment thereof protein is then separated, for example, using two-dimensional gel electrophoresis or size exclusion gel electrophoresis.

In another embodiment, the method of the present invention comprises the additional step of cleavage of an immunoglobulin to facilitate liberation of an immunogenic protein or fragment thereof that is obtained with an immunoglobulin fraction and a fragment of an immunoglobulin. A sample comprising a fragment of an immunoglobulin associated with an immunogenic protein may then be separated using for example, electrophoresis.

In yet another embodiment, an immunoglobulin fraction is obtained and said fraction analysed to identify an immunogenic protein that is obtained with the immunoglobulin

fraction. Preferably, an immunogenic protein is analysed to determine the amino acid sequence of the protein and/or the mass of the protein or a fragment thereof, in order to facilitate identification of the immunogenic protein.

In one embodiment of the invention, an immunoglobulin fraction is obtained from a biological sample. Preferably the biological sample is obtained previously from a subject. In accordance with such an embodiment, the method is performed *ex vivo*.

In yet another embodiment, the subject methods further comprise processing the sample from the subject to produce a derivative or extract that comprises the analyte (eg., protein).

The various embodiments of the invention described herein apply *mutatis mutandis* to a method of identifying a diagnostic, prophylactic or therapeutic target protein. For such target protein identification the subject method preferably comprises:

(i) obtaining an immunoglobulin fraction from a healthy or non-infected subject and from a subject infected by a pathogen;
(ii) identifying proteins or fragments thereof in the immunoglobulin fractions, other than immunoglobulin proteins or fragments thereof; and
(iii) comparing the identified proteins, wherein a protein that is present in the immunoglobulin fraction from the subject infected by the pathogen and not in the immunoglobulin fraction from the healthy subject is a candidate diagnostic target protein or peptide for infection by the pathogen.

The various embodiments of the invention described herein also apply *mutatis mutandis* to a method of identifying a diagnostic, prophylactic or therapeutic target protein. For such target protein identification the subject method preferably comprises:

(i) obtaining an immunoglobulin fraction from a healthy subject or from a subject that does not suffer from an autoimmune disease and from a subject suffering from an autoimmune disease;
(ii) identifying proteins or fragments thereof in the immunoglobulin fractions, other than immunoglobulin proteins or fragments thereof; and
(iii) comparing the identified proteins, wherein a protein that is present in the immunoglobulin fraction from the subject suffering from an autoimmune disease and not in the immunoglobulin fraction from the healthy subject is a candidate diagnostic target protein or peptide for infection by the pathogen.

In another embodiment, a method for target protein identification preferably comprises:

(i) obtaining an immunoglobulin fraction from a healthy subject or from a subject that does not suffer from an autoimmune disease and from a subject suffering from an autoimmune disease;

(ii) identifying proteins or fragments thereof in the immunoglobulin fractions, other than immunoglobulin proteins or fragments thereof; and

(iii) comparing the identified proteins, wherein a protein that is present at an increased level in the immunoglobulin fraction from the subject suffering from an autoimmune disease than the immunoglobulin fraction from the healthy subject is a candidate diagnostic target protein or peptide for infection by the pathogen.

Preferably, the method further comprises confirming the immunogenicity of the identified candidate target protein, such as, for example, by B cell epitope mapping.

A further aspect of the present invention provides a method of producing a diagnostic peptide or protein for infection of a human or other animal by a pathogen, said method comprising identifying a candidate target protein for the infection essentially as described herein and optionally, confirming the immunogenicity of the identified candidate target protein or peptide and/or optionally producing a synthetic or recombinant peptide or protein having an amino acid sequence of an immunogenic protein or peptide present in the immunoglobulin fraction.

<u>Brief description of the figure</u>

Figure 1 is a photographic representation of a 2-dimensional gel showing proteins that have been isolated with an immunological fraction using the method of the present invention.

<u>Detailed description of the preferred embodiments</u>

One aspect of the present invention relates to a method of isolating and/or identifying an immunogenic protein or fragment thereof in a subject capable of eliciting an immune response to the immunogenic protein or peptide, said method comprising:

(i) obtaining an immunoglobulin fraction from the subject; and

(ii) identifying a protein obtained with the immunoglobulin fraction.

As used herein the term "immunogenic protein" shall be understood to mean any peptide, polypeptide or protein that induces an immune response in a subject, such that

a specific antibody is raised against said pathogen derived protein by the subject. In this way, the immunogenic protein will most likely be isolated with an immunoglobulin fraction obtained from a biological sample derived from a subject.

In one embodiment, a method of isolating and/or identifying an immunogenic protein isolated with an immunoglobulin isolates a native or unmodified form of an immunogenic protein. As used herein, the term "native or unmodified form" shall be taken to mean a form of a protein that is produced by the pathogen from which said protein is derived. Accordingly, this term encompasses modifications that are induced by said pathogen or a subject suffering from an autoimmune disease, such as, for example, proteolytic cleavage, post-translational modification, alternative splice forms, and any other modifications result in a change in molecular weight, charge, or amino acid composition.

Following isolation of a protein, the identity of said protein is preferably determined, for example by ascertaining the mass and amino acid sequence of the entire isolated protein or fragment or one or fragments thereof.

In an alternative embodiment, a method of isolating and/or identifying an immunogenic protein isolated with an immunoglobulin isolates a modified form, a fragment or a peptide of an immunogenic protein. As used herein the term "modified form" shall be understood to mean a protein that differs to the native form of said protein. Modifications that may be detected by such methods include, for example, proteolytic cleavage, post-translational modification, and any other modifications result in a change in molecular weight, charge, or amino acid composition. In accordance with this embodiment, a modified form, a fragment or a peptide of an immunogenic protein may be produced as a result of an immune response by a subject from whom a biological sample is derived. During such an immune response, many proteolytic enzymes, such as, for example neutrophil elastase and pathogen derived elastases are produced, often resulting in the cleavage and degradation of a protein.

As used herein the term "immunoglobulin fraction" shall be taken to mean a component of a biological sample that is isolated with an immunoglobulin. Such a fraction may comprise, for example, an immune complex, an antibody-HLA complex, immunoglobulin light chain, immunoglobulin heavy chain, a component of the

complement pathway, fibrinogen, haptoglobin and serum albumin amongst other components.

As used herein the term "immunoglobulin" shall be taken to mean a protein produced by lymphocytes, where said protein preferably has specific antibody activity. Preferably an immunoglobulin comprises four polypeptide chains, two identical heavy chains and two identical light chains, linked by disulphide bonds. It is preferred that an immunoglobulin is selected from the group consisting of IgA, IgD, IgE, IgG, and IgM.

As used herein the term "antibody" refers to intact monoclonal or polyclonal antibodies, immunoglobulin (IgA, IgD, IgG, IgM, IgE) fractions, humanized antibodies, or recombinant single chain antibodies, as well as fragments thereof, such as, for example Fab, F(ab)2, and Fv fragments.

In accordance with this aspect of the invention, an entire immunoglobulin fraction is obtained from a subject. Accordingly, this process is performed without purifying free immunoglobulin from immunoglobulin that is complexed with an antigen or immunogenic protein.

In a particularly preferred embodiment, an immunoglobulin fraction is an immunoglobulin G (IgG) fraction.

In one embodiment, an immunogenic protein or peptide is a pathogen derived protein.

As used herein the term "pathogen" shall be taken in its broadest context to include any organism, microorganism or particle that is able to infect another organism, preferably resulting in disease and/or inflammation. Such a pathogen may infect any Metazoan, or any member of the kingdom Animalia. Accordingly, such a pathogen may be any pathogen, for example a viral pathogen, a bacterial pathogen, a fungal pathogen, a pathogen of the phylum nematoda, that infects any animal.

For example a pathogen is a pathogen selected from the group consisting of *Actinobacillus* spp., *Aeromonas spp.*, *Actinomyces* spp., *Lactobacillus* spp., *Eubacterium* spp., *Bifidobacterium* spp., *Bacteroides spp.*, *Bacillus spp.*, *Bacteroides spp.*, *Bordetella spp.*, *Borrelia spp.*, *Brucella spp.*, *Burkholderia spp.*, *Campylobacter spp.*, *Clostridium spp.*, *Corynebacterium spp.*, *Enterobacter* spp., *Klebsiella* spp.,

Salmonella spp., *Serratia* spp., *Shigella spp.*, *Yersinia spp.*, *Enterococcus spp.*, *Escherichia coli*, *Haemophilus spp.*, *Helicobacter spp.*, *Klebsiella pneumoniae* *Legionella pneumophila*, *Leptospira interrogans*, *Listeria monocytogenes*, *Micrococcus spp.*, *Moraxella catarrhalis* ,*Morganella* spp., *Mycobacterium spp.*, *Mycoplasma* spp. *Nocardia spp.* *Neisseria spp.*, *Pasteurella multocida*, *Plesiomonas shigelloides*, *Propionibacterium acnes*, *Proteus* spp., *Pseudomonas aeruginosa*, *Salmonella spp.*, *Shigella spp.*, *Staphylococcus spp.*, *Streptococcus spp.*, *Streptomyces spp.*, *Treponema spp.*, *Vibrio spp.*, *Yersinia spp.*, *Chlamydia spp*, *Clostridium spp*, *Listeria spp*, *Plasmodium spp*, *Tryponosoma spp*, and a virus selected from the group consisting of, a virus from the family Astroviridae, Caliciviridae, Picornaviridae, Togaviridae, Flaviviridae, Caronaviridae, Paramyxviridae, Orthomyxoviridae, Bunyaviridae, Arenaviridae, Rhabdoviridae, Filoviridae, Reoviridae, Bornaviridae, Retroviridae, Poxviridae, Herpesviridae, Adenoviridae, Papovaviridae, Parvoviridae, Hepadnaviridae. As will be apparent to the skilled artisan this is not an exhaustive list, and this invention relates to the isolation of an immunogenic protein from any pathogen from a biological sample derived from a subject.

In another embodiment, an immunogenic protein or peptide is a protein expressed or produced by the subject from whom the immunoglobulin fraction is obtained. For example, such a protein may be associated with an autoimmune disease, such as, for example, rhematoid arthritis, multiple sclerosis, type-1 diabetes, inflammatory bowel disease, Crohn's Disease, ulcerative colitis, systemic lupus erythemotasus, psoriasis, scleroderma, autoimmune thyroid disease, central nervous system vasculitis, and autoimmune myositis.

In a particularly preferred embodiment, an immunogenic protein or peptide is bound by an immunoglobulin, facilitating isolation of said an immunogenic protein or peptide with an immunoglobulin. Accordingly, it is likely that an immunogenic protein or fragment thereof has elicited an immune response in a subject such that a B-lymphocyte (or cell) produces IgD and IgM that specifically bind to the immunogenic protein. It is particularly preferred that said B-lymphocyte differentiates into a plasma cell secretes IgM, IgG, IgE and/ or IgA that specifically bind to the immunogenic protein. Consequently, isolation of an immunoglobulin fraction facilitates both isolation and identification of the immunogenic protein or fragment thereof.

]

Obtaining an immunoglobulin fraction

Preferably an immunoglobulin fraction is obtained from a biological sample that is derived from a subject. More preferably, an immunoglobulin is obtained from a biological sample selected from the group consisting of, whole blood, plasma, serum, sputum, saliva, pleural fluid, pericardial fluid, peritoneal fluid, lymph fluid, a lymph node, a spleen or fragment thereof, a peripheral blood mononuclear cell, a lymphocyte, a B-lymphocyte, a T-lymphocyte, a helper T-cell, a cytotoxic T-cell, a macrophage, a dendritic cell, a polymorphonuclear cell, and a mast cell.

In a particularly preferred embodiment, a biological sample is derived previously from a subject using a method well known in the art, for example using a syringe.

In one embodiment, an immunoglobulin is isolated from a biological sample by precipitatation using, for example, ethanol, polyethylene glycol, lyotropic (anti-chaotropic) salts such as ammonium sulfate and potassium phosphate.

In another embodiment, an immunoglobulin is isolated from a biological sample by ion exchange chromatography, essentially as described in Burnouf and Radosevich, *J Biochem Biophys Methods, 49(1-3)*, 575-86, 2001.

In yet another embodiment, an immunoglobulin is isolated from a biological sample by hydrophobic chromatography, essentially as described in Doellgast and Plout, *Immunochemistry, 13(2)*, 135-139, 1976. Such a method utilises a matrix that binds an immunoglobulin in the presence of lyotropic salts. Accordingly, lyotropic salts are added to a biological sample derived from a subject and this sample contacted to said matrix. An immunoglobulin is then released from said matrix by reducing the concentration of lyotropic salts in a sample in a stepwise manner.

In a further embodiment, an immunoglobulin is isolated using thiophilic adsorption chromatography, essentially as described in Porath *et al, FEBS Letters, 185*, 306, 1985 and Knudsen *et al, Analytical Biochemistry, 201*, 170, 1992. This method comprises the use of divinyl sulfone activated agarose that to which has been bound to various ligands comprising a free mercapto- group. These ligands specifically bind an immunoglobulin in the presence of potassium sulfate (0.5M). Such ligands include, for example 2-mercaptopyridine, 2-mercaptopyrimidine, and 2-mercaptothiazoline. Again,

an immunoglobulin is released from a ligand by reducing the concentration of a lyotropic salt (ie potassium sulfate) in a sample.

In yet another embodiment, an immunoglobulin is isolated using a matrix that specifically binds to an immunoglobulin, such as, a matrix described in US Patent No. 6,498,016. Such a matrix comprises a solid phase backbone, such as, for example cellulose, agarose, dextran based beads or organic polymers; optionally a spacer element; and a ligand comprising an aromatic or a heteroaromatic group, preferably a benzene ring fused with a heteroaromatic ring system. Such a matrix does not require the use of a lyotropic salt, rather it is able to bind an immunoglobulin under neutral conditions. An immunoglobulin is eluted or dissociated from such a matrix using conditions well known in the art, such as, for example, washing the matrix with a buffer with a reduced pH, for example glycine, pH3.

In an alternative embodiment, an immunoglobulin fraction is separated from other constituents by affinity chromatography on Kaptive-M™-Sepharose. Those skilled in the art are aware that IgM binds to Kaptiv-M, the active constituent of which is a peptidomimetic. Accordingly, a fraction comprising IgM and an immunogenic protein bound by IgM is isolated using this method.

Alternatively, MBP-Sepharose is used. Those skilled in the art are aware that MBP binds to mannose residues present on the IgM Fc5μ region, and, as a consequence, is specific for IgM. The initial step of binding is performed under native conditions so as not to perturb any protein-protein interaction (e.g., MBP bound to antibody, or alternatively, an antibody-antibody interaction). IgM and an antigenic protein are eluted from the affinity matrix using a dissociating buffer, such as, for example, a buffer comprising a high salt concentration (e.g., 3M $MgCl_2$ in HEPES pH 7.2) that releases the antibodies and immunogenic proteins as unbound components.

In a preferred embodiment, an immunoglobulin is isolated using Protein G, Protein A or Protein L. Methods of isolation of an immunoglobulin using protein G are well known in the art and are described, for example, in Bjorck and Kronvall, *J. Immunol. 33(2)*, 969-974, 1984. Methods of isolation of an immunoglobulin using protein A are well known in the art and are described, for example, in Hjelm *et al*, *FEBS Lett 28(1)*

73-76 1972. Methods of isolating an immunoglobulin using protein L are well known in the art and are described, for example, in Akerstrom and Bjorck *J Biol Chem 264(33)* 19740-19746, 1989.

As will be known to the skilled artisan, protein A, protein G and protein L bind to an immunoglobulin fraction. Accordingly, an immunoglobulin fraction is amenable to purification by affinity chromatography on Protein-A Sepharose or Protein-G Sepharose or Protein L Sepharose. The initial step of binding is performed under native conditions so as not to perturb any protein-protein interaction (e.g., Protein-A or Protein-G bound to antibody, which is in turn bound to an immunogenic protein. Antibodies are eluted from the protein-A, protein-G or protein-L using a dissociating buffer, such as, for example, a buffer comprising a high salt concentration (e.g., 3M $MgCl_2$ in HEPES pH 7.2) that releases the antibodies as unbound components.

In a particularly preferred embodiment, an immunoglobulin fraction is isolated using protein G.

As used herein, the term "protein G" shall be taken to include a protein comprising one or more natural IgG-binding domains of protein G, a hybrid or fusion protein comprising an IgG-binding domain of a native or naturally-occurring protein G, or a mutant or variant of a native or naturally-occurring protein G that retains the ability of native protein G to bind IgG, or a fragment of a native or naturally-occurring protein G that retains the ability of native protein G to bind IgG.

As used herein, the term "protein A" shall be taken to include a protein comprising one or more natural IgG-binding domains of protein A, a hybrid or fusion protein comprising an IgG-binding domain of a native or naturally-occurring protein A, or a mutant or variant of a native or naturally-occurring protein A that retains the ability of native protein A to bind IgG, or a fragment of a native or naturally-occurring protein A that retains the ability of native protein A to bind IgG.

As used herein, the term "protein L" shall be taken to include a protein comprising one or more natural antibody light-chain-binding domains of protein L, a hybrid or fusion protein comprising an antibody light-chain-binding domain of a native or naturally-occurring protein L, or a mutant or variant of a native or naturally-occurring protein L that retains the ability of native protein L to bind an antibody light-chain, or a fragment of a native or naturally-occurring protein L that retains the ability of native protein L to bind an antibody light-chain.

Protein G, protein A or protein L referred to herein are obtained from a commercial source, or alternatively, produced by conventional means. Commercial sources will be well known to those skilled in the art.

For example, protein G, protein A and/or protein L are available from Amersham-Pharmacia, Castle Hill, NSW, Australia.

Alternatively, protein G, protein A or protein L are isolated using the methods described in US Patent No. 4,945,157, US Patent No. 6,555, 661 and US Patent No. 4,876,194 respectively. Recombinant protein G, protein A or protein L may also be produced using techniques well known in the art, as described, for example, in Sambrook *et al* (*In:* Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, New York, Second Edition (1989), whole of Vols I, II, and III). For example recombinant protein G may be produced using the methods described in US Patent No. 5,082,773.

As will be apparent to the skilled artisan, protein G, protein A and/or protein L may be attached to a solid support in order to allow for affinity purification of an immunoglobulin. Solid phase supports suitable for attachment of protein G, protein A and/or protein L include a solid phase support selected from the group consisting of polymers having hydroxyl groups, either free or in esterified form, such as agarose, cellulose, including cellulose esters such as cellulose nitrate, diazocellulose, cellulose acetate, cellulose propionate, and acrylamide polymers and copolymers, such as polyacrylamide and acrylamide, microtitre plates, glass, polystyrene, polypropylene, polyethylene, dextran, nylon, agar, starch, and a chemically active membrane having a large surface area comprising a hydrophobic, microporous, skinless, polyamide

membrane which is chemically bound to a residue of an activating agent which is capable of immobilizing a. protein G, protein A and/or protein L.

Protein G, protein A and/or protein L may be immobilized to a solid phase support according to methods known to those of ordinary skill in the art. For example, a protein G, protein A and/or protein L is coated or bonded, either covalently or by adsorption, to a solid phase. Methods for immobilizing a protein to a solid phase support are taught, for example, in U.S. Pat. No. 3,652,761, U.S. Pat. No. 3,879,262, U.S. Pat. No. 3,986,217, and U.S. Pat. No. 4,693,985. Preferably, a protein G, protein A and/or protein L are immobilized to tresyl activated or cyanogen bromide activated agarose or a maleimide-activated agarose support. This support may be prepared by treating agarose modified to contain primary amino groups (in particular, AH-Sepharose, Pharmacia Co.) with sulfosuccinimidyl-4-(p-maleimidophenyl) butyrate.

Immobilised protein G, protein A and/or protein L is of particular use in affinity purification of an immunoglobulin. Affinity purification techniques are well known in the art and are described, for example, in Scopes (*In:* Protein purification: principles and practice, Third Edition, Springer Verlag, 1994). Methods of affinity purification typically involve contacting a biological sample isolated from a subject to an immobilised Protein G, Protein A and/or Protein L, and, following washing, eluting an immunoglobulin that is bound to a Protein G, Protein A and/or Protein L.

Alternatively, a protein G, protein A and/or protein L is covalently bound to a molecule, such as, for example, biotin. Accordingly, an affinity purification method involving such a conjugated protein G, protein A and/or protein L, uses, for example, streptavidin that has been conjugated to a solid support to bind, or capture a conjugated protein G, protein A and/or protein L.

In another embodiment, a protein G, protein A and/or protein L is covalently linked to a magnetic or paramagnetic bead, such as for example a Dynabead® (available from Dynal Biotech, Oslo, Norway). Accordingly, such a bead is contacted with a biological sample derived from a subject. Said bead and any bound immunoglobulin is then isolated from a subject using a magnetic or paramagnetic field.

In yet another embodiment, protein G, protein A and/or protein L is contacted with a biological sample derived from a subject for a time and under conditions that allow said protein G, protein A or protein L to bind to an immunoglobulin fraction. This sample is then bound by an antibody that specifically recognises a protein G, protein A or protein L. Preferably, this antibody is bound to a solid support or some other means that allows for isolation of a protein G, protein A and/or protein L from a biological sample, such as for example, agarose, a plastic solid support or a glass solid support. Polyclonal antibodies that specifically recognise protein G, protein A or protein L are available from Sapphire Bioscience, Crows Nest, NSW, Australia.

In yet another embodiment, in an additional step, an immunoglobulin that is bound to an immunogenic protein is covalently bound by hydrazide, facilitating liberation of an immunogenic protein from a bound immunoglobulin. Such a method is performed essentially as described in O'Shannessy and Hoffmann, *Biotechnol. Appl. Biochem. 9(6)*, 488-496, 1987. Essentially this method consists of isolating an immunoglobulin using a method known in the art and/or described herein, and oxidising said immunoglobulin with sodium periodate. This oxidisation results in the formation of aldehydes on any oligosaccharide moiety. An oxidised sample is then contacted with a hydrazide-derivatized solid support. This results in the formation of a stable hydrazone linkage between an oxidised immunoglobulin and said solid support. Suitable solid supports include, for example, agarose, a glass bead or a polystyrene, polypropylene or polycarbonate bead or microtitre plate. As a hydrazone bond is extremely stable, even at low pH ranges, an immunogenic protein that is bound to an immunoglobulin is dissociated by a method known in the art and/or described herein. Accordingly, an immunogenic protein is dissociated from an immunoglobulin, facilitating further analysis of this protein without any contamination by an immunoglobulin.

In one embodiment, such a protein sample is directly analysed to determine the identity of an immunogenic protein using methods known in the art and/or described herein.

In another embodiment, a protein sample isolated essentially using the method of O'Shannessy and Hoffmann *supra* is separated using a method known in the art and/or described herein prior to analysis to determine the identity of an immunogenic protein.

<u>Dissociating an immunogenic protein from an immunoglobulin</u>

In one embodiment an immunogenic protein is dissociated from an immunoglobulin to which it is bound, prior to further analysis. An immunogenic protein is considered to be dissociated from an immunoglobulin, when it is no longer bound by said immunoglobulin, that is the immunogenic protein and the immunoglobulin are separate from one another. Methods of liberating a protein from an immunoglobulin are well known in the art and are described, for example, in Scopes (*In:* Protein purification: principles and practice, Third Edition, Springer Verlag, 1994). For example, an immunogenic protein is liberated from an immunoglobulin by changing the pH of a sample (for example with glycine (pH 3) or triethanolamine (pH 11)), increasing the salt concentration of a sample (for example with 5M Lithium Chloride), treating a sample with an ionic detergent (for example sodium-doedecyl sulfate (SDS)), treating a sample with a dissociating agent (for example urea) or treating a sample with a chaotropic agent (for example thiocyanate). As will be apparent to a skilled artisan an immunogenic protein may be liberated from an immunoglobulin using a combination of the previously mentioned methods.

In a preferred embodiment, an immunogenic protein is dissociated from an immunoglobulin by changing the pH of a sample with glycine. Preferably, the glycine is at a pH of about 1.5 to a pH of about 4, more preferably, a pH of about 1.9 to a pH of about 2.7 and most preferably a pH of about 2.3 to a pH of about 2.7.

In an alternative embodiment, an immunogenic protein is isolated from an immunoglobulin fraction using caproylic acid and ammonium sulphate precipitation. As will be known to the skilled artisan, these agents provide preparations that comprise essentially an immunoglobulin or an immunogenic protein. By dissolving the immunoglobulin fraction in a dissociating buffer such as, for example, a high-salt buffer (e.g., 3M $MgCl_2$ in / HEPES pH 7.2), the immunogenic protein is released as unbound components. The immunogenic protein fraction and the immunoglobulin fraction are then separated by SEC, for example, using the dissociating buffer as an eluant to maintain the IgG and IgM components as unbound components.

In an alternative embodiment, an immunoglobulin fraction is subjected to free-flow electrophoresis under denaturing conditions. A biological sample is clarified and the

proteins are precipitated under conditions that leave immunoglobulin and some α_2-macroglobulins in solution. The immunoglobulin fraction is then precipitated and redissolved in a suitable buffer, applied to a free-flow electrophoresis (FFE) device (e.g., Octopus™, Tecan™) for separation by continuous solution-phase isoelectric focusing as described essentially by Hoffman *et al., Proteomics 1*, 807-818, 2001). Fractions are obtained, preferably corresponding to an immunogenic protein, and exchanged into a suitable buffer (e.g., PBS) using PD-10 or fast–desalting columns (Amersham Biosciences) prior to further analysis to determine the identity of the immunogenic protein using methods well known in the art and/or described herein.

Separation of an immunogenic protein and an immunoglobulin fraction

In a preferred embodiment, following dissociation of an immunogenic protein from an immunoglobulin, an immunogenic protein is isolated or separated from an immunoglobulin. Methods of separating proteins are well known in the art and are described, for example, in Scopes (*In:* Protein purification: principles and practice, Third Edition, Springer Verlag, 1994).

In another preferred embodiment, an immunoglobulin fraction is separated without first dissociating an immunoglobulin and an immunogenic protein.

In one embodiment, an immunogenic protein is separated from an immunoglobulin following dissociation using native gel electrophoresis. As used herein the term "native gel electrophoresis" shall be taken to mean any form of electrophoresis that is performed under conditions that do not denature a protein, that is a protein that is electrophoresed retains its native size, shape and charge. Accordingly, mobility of a protein using native gel electrophoresis depends upon both the charge of the protein and the hydrodynamic size of the protein. Such a method is of particular use in the separation of an immunogenic protein and an immunoglobulin, as, not only does native gel electrophoresis maintain the size, shape and charge of a protein, this method also allows proteins that normally interact to remain bound. As an immunoglobulin comprises two heavy chains and two light chains, it is expected that an immunoglobulin has a molecular weight of at least about 150 kDa (corresponding to the predicted molecular weight of IgG). Accordingly, electrophoresis of a sample

prepared using the previously described method facilitates separation of an immunological protein from an immunoglobulin.

In one embodiment, a sample comprising an immunogenic protein and an immunoglobulin are electrophoresed using 1 dimensional native gel electrophoresis using techniques well known in the art. In such cases proteins are merely separated by their molecular weight and charge. Accordingly, such a method is of use in separating an immunoglobulin from a smaller immunogenic protein. The immunogenic protein is then isolated and/or identified using methods well known in the art and/or described herein.

In another embodiment, a sample comprising an immunogenic protein and an immunoglobulin are electrophoresed using native two-dimensional gel electrophoresis. For example proteins are separated in one dimension using isoelectric focussing. Using such a method, proteins are separated by their isoelectric point, that is the pH at which the net charge of a protein is equal to zero. In order to separate proteins by their isoelectric point a sample is electrophoresed in a gel that comprises a pH gradient. Under such conditions, a protein will move to a position on said gradient where its net charge is equal to zero. Following isoelectric focussing proteins are separated according to their mass, using standard native gel electrophoresis. Accordingly, such a method is of use in the separation of an immunoglobulin from an immunogenic protein.

In yet another embodiment, an immunogenic protein is isolated from an immunoglobulin by dissociating said immunogenic protein from an immunoglobulin using a method known in the art and/or described herein and separating said pathogen derived protein and said immunoglobulin using a gel filtration column. Such columns are available from commercial sources, such as, for example, Sigma-Aldrich or Amersham-Pharmacia. Methods of gel filtration are well known in the art and are described, for example, in Scopes (*In:* Protein purification: principles and practice, Third Edition, Springer Verlag, 1994). Gel filtration chromatography separates proteins based upon their size. Such methods comprise contacting a sample to a column that comprises a solid matrix that consists of a specified pore size. Proteins that are of a sufficiently low molecular weight fit within these pores and are said to be included, while those that do not are excluded. Proteins are eluted from said column, with those that are excluded eluting prior to those that are included. Accordingly, as an immunoglobulin in its native state is a relatively large molecule, it is most likely that

said immunoglobulin will elute before an immunogenic protein. Following collection of a sample comprising a immunogenic protein, said sample is analysed using a method well known in the art, and described herein. Alternatively, a sample comprising an immunogenic protein is separated using electrophoresis, for example native or denaturing one- or two-dimensional gel electrophoresis, prior to any analysis of said immunogenic protein.

In a further embodiment, an immunogenic protein is isolated from an immunoglobulin following dissociation from said immunoglobulin using other methods of size exclusion, such as for example, centrifugation using a size exclusion filter (for example as available from Millipore), high performance liquid chromatography or reverse phase chromatography, amongst others.

In a further embodiment, an immunogenic protein is separated from an immunoglobulin following dissociation from said immunoglobulin using density gradient fractionation. Methods of fractionation using a density gradient are well known in the art. For example, proteins may be separated using ultracentrifugation, where a sample is added to a linear sucrose gradient ranging, for example, from 5% to 20% and subsequent centrifugation. Accordingly, proteins are separated with regard to centrifugal force, frictional force and boyant force. Using such a method, an immunoglobulin is separated from an immunogenic protein as it is a relatively large protein. Following separation from an immunoglobulin, an immunogenic protein is analysed using a method well known in the art and/or described herein.

In a particularly preferred embodiment, an immunogenic derived protein is isolated from an immunoglobulin following dissociation from said immunoglobulin using denaturing electrophoresis. Alternatively, a sample that has not been previously dissociated is separated using denaturing electrophoresis. Denaturing electrophoresis is performed as described herein, however, rather than being performed under native conditions, reagents that denature proteins are included in either or both the electrophoresis gel and in sample preparation. Accordingly, protein samples are denatured using, for example, detergent (eg SDS), or other denaturants (eg 2-mercaptoethanol, DTT and/or heat).

In one embodiment, an immunogenic protein is isolated from an immunoglobulin using reducing one-dimensional gel electrophoresis, using methods known in the art, and

described, for example, in Scopes (*In:* Protein purification: principles and practice, Third Edition, Springer Verlag, 1994). In accordance with this embodiment, proteins are separated by their molecular weight. Accordingly, only a protein that has a molecular weight different to both a heavy and a light chain of an immunoglobulin is readily detectable using this method.

In another embodiment, an immunogenic protein is isolated from an immunoglobulin using reducing two-dimensional gel electrophoresis, as is known in the art. In accordance with this embodiment, proteins are separated, for example, by their isoelectric point or net charge and molecular weight. As such, this method is of particular use in determining a protein that has a different molecular weight and/or isoelectric point from that of an immunoglobulin light or heavy chain.

In accordance with either of the two previous embodiments, following separation of an immunogenic protein from an immunoglobulin using reducing electrophoresis, an immunogenic protein is identified using methods well known in the art and/or described herein.

In one embodiment, sample comprising an immunogenic protein is separated using one dimensional or two dimensional electrophoresis prior to any further analysis.

In accordance with any of the previously mentioned embodiments relating to dissociation or separation of an immunogenic protein from an immunoglobulin, a sample comprising an isolated pathogen derived protein may optionally be concentrated prior to further analysis. Methods of concentrating a protein are known to those skilled in the art, and include, for example, precipitation, freeze drying, use of funnel tube gels (TerBush and Novick, *Journal of Biomolecular Techniques, 10(3)*; 1999), ultrafiltration or dialysis.

<u>Identification of an immunogenic protein without dissociation from an immunoglobulin</u>

In yet another embodiment, an immunogenic protein is not dissociated from an immunoglobulin. An immunoglobulin fraction is isolated from a biological sample derived from a subject using a method well known in the art and/or described herein, and the Fc region of said immunoglobulin is cleaved using a protease. A particularly preferred protease is a protease selected from the group consisting of, papain, elastase, SpeB and EndoS from *Streptococcus pyogenes* and pepsin. Both papain and elastase

are commercially available from, for example, Merck, while pepsin is commercially available from Calzyme Laboratories, San Luis Obispo, CA, USA.

In accordance with this embodiment, an immunogenic protein remains bound to a fragment of an immunoglobulin, and said complex is liberated from the Fc region of said immunoglobulin. This sample is then analysed using any method known in the art and/or described herein.

Alternatively, an immunogenic protein and a cleaved immunoglobulin are first separated using methods well known in the art and/or described herein, such as, for example, non-reducing or reducing one- or two-dimensional gel electrophoresis, prior to analysis to determine the identity of the immunogenic protein.

In another embodiment, an immunogenic protein is obtained from a biological sample using a protein chip. To produce such a protein chip, a protein that is able to bind an immunoglobulin of interest, such as, for example protein G, protein A and/or protein L is bound to a solid support such as for example glass, polycarbonate, polytetrafluoroethylene, polystyrene, silicon oxide, metal or silicon nitride. This immobilization is either direct (e.g. by covalent linkage, such as, for example, Schiff's base formation, disulfide linkage, or amide or urea bond formation) or indirect. Methods of generating a protein chip are known in the art and are described in for example U.S. Patent Application No. 20020136821, 20020192654, 20020102617 and U.S. Patent No. 6,391,625. In order to bind a protein to a solid support it is often necessary to treat the solid support so as to create chemically reactive groups on the surface, such as, for example, with an aldehyde-containing silane reagent. Alternatively, a protein that is able to bind an immunoglobulin of interest may be captured on a microfabricated polyacrylamide gel pad and accelerated into the gel using microelectrophoresis as described in, Arenkov *et al. Anal. Biochem. 278*:123-131, 2000.

A biological sample comprising an immunoglobulin fraction is then contacted to the protein chip and the immunoglobulin fraction isolated from said biological sample.

In one embodiment, biomolecular interaction analysis-mass spectrometry (BIA-MS) is used to rapidly detect and characterise an immunogenic protein bound to said immunoglobulin (Nelson *et al. Electrophoresis 21*: 1155-1163, 2000). One technique

useful in the analysis of a protein chip is surface enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS) technology to characterise a protein bound to the protein chip. Alternatively, the protein chip is analysed using ESI as described in U.S. Patent Application 20020139751.

Analysis of an immunogenic protein

Following isolation and/or dissociation of an immunogenic protein from an immunoglobulin fraction, or cleavage of an immunoglobulin and recovery of an immunogenic protein, it is particularly preferred that a sample comprising said immunogenic protein be analysed in order to establish the identity of said protein. Methods of analysis of a protein in order to determine the identity of said protein are well known in the art and include, a method selected from the group consisting of Edman sequencing, mixed peptide sequencing, mass spectrometry including MALDI, TOF, ESI and ion trap analysis.

In one embodiment, the identity of an immunogenic protien is identified using Edman sequencing (as described by Edman, *Arch. Biochem. Biophys.*, *22*, 475-483, 1949) in order to determine the N-terminal sequence of an immunogenic protien and comparing this sequence to a known sequence in order to determine the identity of said pathogen derived protein. In accordance with this embodiment, it is particularly preferred that an immunogenic protien is separated from a contaminating molecule, such as, for example another protein, prior to Edman sequencing. Preferably, an immunogenic protein is isolated from a gel prior to sequencing, using methods well known in the art. Following isolation of an immunogenic protein, the amino terminus of said protein is derivatized with phenylisothiocyanate under basic conditions. Preferably the base used in this step is a non-nucleophile such as, for example, a triethylamine or diisoproylethylamine. This coupling step produces a phenylthiocarbamyl peptide or protein. The thiocarbonyl function of the phenylthiocarbamyl peptide or protein is a moderately strong nucleophile, and under acidic conditions it will cleave the carbonyl carbon of the adjacent peptide bond. This cleavage step results in the production of an anilothiazolinone of the terminal amino acid and leaves the original peptide or protein shortened by exactly one amino acid residue. The anilothiazolinone of terminal amino acid has different solubility properties from the peptide or protein, so it can be extracted and subjected to further analysis. The shortened peptide or protein again has a bare amino terminus, so it can be subjected to additional cycles of coupling, cleavage, and extraction.

The extracted anilothiazolinone of the terminal amino acid, however, is not stable. Under acidic aqueous conditions, anilothiazolinones rearrange rapidly to form more stable phenylthiohydantoins, which are more amenable to analysis. A stable phenylthiohydantoin is then analyzed by UV absorbtion detection reverse phase high performance liquid chromatography, to determine the identity of the terminal amino acid.

Following determining the N-terminal sequence of an immunogenic protein, this sequence is compared to a database of sequences in order to determine whether or not the derived sequence is identical to or homologous to a known sequence. Such a database is available, for example, at NCBI.

As used herein the term "NCBI" shall be taken to mean the database of the National Center for Biotechnology Information at the National Library of Medicine at the National Institutes of Health of the Government of the United States of America, Bethesda, MD, 20894.

In determining whether or not two amino acid sequences fall within the defined percentage identity limits supra, those skilled in the art will be aware that it is possible to conduct a side-by-side comparison of the amino acid sequences. In such comparisons or alignments, differences will arise in the positioning of non-identical residues depending upon the algorithm used to perform the alignment. In the present context, references to percentage identities and similarities between two or more amino acid sequences shall be taken to refer to the number of identical and similar residues respectively, between said sequences as determined using any standard algorithm known to those skilled in the art. In particular, amino acid identities and similarities are calculated using software of the Computer Genetics Group, Inc., University Research Park, Maddison, Wisconsin, United States of America, eg., using the GAP program of Devereaux *et al.*, *Nucl. Acids Res. 12*, 387-395, 1984, which utilizes the algorithm of Needleman and Wunsch, *J. Mol. Biol. 48*, 443-453, 1970. Alternatively, the CLUSTAL W algorithm of Thompson *et al.*, *Nucl. Acids Res. 22*, 4673-4680, 1994, is used to obtain an alignment of multiple sequences, wherein it is necessary or desirable to maximise the number of identical/similar residues and to minimise the number and/or length of sequence gaps in the alignment. Amino acid sequence alignments can

also be performed using a variety of other commercially available sequence analysis programs, such as, for example, the BLAST program available at NCBI.

In another embodiment, an immunogenic protein is identified using mixed-peptide sequencing, as described in Damer *et al, J. Biol. Chem. 273*, 24396-24405, 1998. In accordance with this embodiment, an immunogenic protein is cleaved into peptides using cyanogen bromide or skatole and these peptides are sequenced using the Edman sequencing method, simultaneously.

In another embodiment, an immunogenic protein is identified using mass spectrometry. In accordance with this embodiment an immunogenic protein is separated using electrophoresis in order to remove a contaminating particle, such as for example another protein. Preferably, an immunogenic protein is digested with a protease prior to analysis with mass spectrometry.

In one embodiment, following separation of an immunogenic protein using electrophoresis, it is preferred that said protein is digested in the gel in which electrophoresis occurred. In-gel digestion of a protein, peptide or polypeptide enables more of said protein to be recovered from a gel than other methods such as for example electroblotting. Accordingly, the increased quantity of an immunogenic protein facilitates easier analysis of said protein. Methods of in-gel digestion are well known in the art and are described, for example, in Schevenko *et al, Anal. Chem., 68*, 850-858, 1997. Furthermore, kits that facilitate in-gel digestion of a protein are commercially available, for example, from Millipore, Billerica, MA 01821, USA.

In one embodiment an immunogenic protein or a peptide thereof is purified and optionally concentrated prior to further analysis. Preferably an immunogenic protein or a peptide thereof is purified using reverse-phase chromatography.

In an alternate embodiment, an immunogenic protein is not electrophoresed, rather a sample dissociated from an immunoglobulin is used for analysis. Optionally, an immunogenic protein is digested with a protease, such as, for example, trypsin, in order to facilitate analysis of peptides of an immunogenic protein. Accordingly, following purification, and optionally concentration, such a sample is analysed by mass spectrometry.

Following purification of an immunogenic protein or a peptide thereof, samples are ionised.

In one embodiment, a sample is ionised using electrospray ionisation (ESI), as described by, for example Fenn *et al*, *Science*, *246*, 64-71, 1989 and Wilm *et al*, *Nature*, *379*, 466-469, 1996. In accordance with this embodiment a sample comprising an immunogenic protein or a fragment thereof is forced into a mass spectrometer through a microcapillary tube. A potential difference between the chamber of the mass spectrometer and the microcapillary tube cause the sample comprising an immunogenic protein or a fragment thereof to be ejected from said tube as a fine mist. As the liquid in this mist evaporates (ie the solution in which a protein is suspended) the protein or peptide thereof becomes desolvated. Accordingly, a protein or peptide is converted to ions.

In another embodiment, a sample is ionised using matrix assisted laser desorption/ionisation (MALDI), as described by, for example, Karas and Hillenkamp, *Anal. Chem.*, *60*, 2299-2301, 1988. In accordance with this embodiment, a sample is incorporated into a matrix, such as for example a-cyano-4-hydroxycinnamic acid, 3,5 dimethoxy-4-hydroxycinnamic acid (Sinapinic acid) or 2,5 dihydroxybenzoic acid (Gentisic acid). The sample and matrix are then spotted onto a metal plate and subjected to irradiation by a laser. This promotes the formation of molecular ions. As will be apparent to those skilled in the art, variations of this method are clearly encompassed in the instant invention, such as, for example, atmospheric pressure MALDI.

As will be apparent to the skilled artisan other forms of ionization are clearly encompassed in the instant invention, for example, atmospheric pressure chemical ionization.

Following ionisation of a protein or peptide thereof the mass of these molecular ions are analysed.

In one embodiment, the mass of a molecular ion is analysed using a quadrupole mass analyser, or a linear quadripole, essentially as described in Burlingame *et al*, *Anal. Chem.* *70*, 674R-716R and references cited therein. In accordance with this embodiment, an ion is transmitted through an electric field generated by an array of

four metallic rods, to which rf and dc voltages are supplied. This voltage causes an ion to oscillate with the frequency of this oscillation depending upon the m/z value of the ion. Only those ions that show a stable oscillation, that is those that have a given m/z value as determined by the rod assembly, oscillation frequency, rf voltages and dc voltage, are retained for further analysis. Accordingly this facilitates analysis of the mass to charge ratio (m/z) of a peptide or protein. This is then compared to a library of molecular weights, such as, for example using database search software provided by the UK Human Genome Mapping Project Resource Centre.

Though combination of multiple quadripoles the amino acid sequence of a protein or peptide is determined.

In another embodiment, the mass of a molecular ion is analysed using an ion trap mass analyzer, essentially as described in Cooks *et al, Chem. Eng. News, 69*, 26, 1991. This form of analysis is a form of a quadrupole mass analyser where the generators of an electric charge are arrayed in three dimensions rather than in a linear fashion. In accordance with this embodiment, a molecular ion of a m/z ratio is trapped in a three-dimensional electric field. An ion trap mass analyzer is also of particular use in tandem mass spectrometry (MS/MS) experiments for the determination of a sequence of a peptide, polypeptide or protein. Methods of MS/MS are known in the art and/or described herein.

In yet another embodiment, the mass of a molecular ion is analysed by its time of flight (TOF), essentially as described by Yates, *J. Mass Spectrom. 33*, 1-19, 1998 and references cited therein A time of flight instrument measures the m/z ratio of an ion by determining the time required for it to traverse the length of a flight tube. Optionally, such a TOF mass analyzer includes an ion mirror at one end of the flight tube in order to reflect said ion back through the flight tube to a detector. Accordingly, an ion mirror serves to increase the length of a flight tube, increasing the accuracy of this form of analysis.

Time of flight analysis is also useful for determining the mass and therefore the predicted sequence of a peptide, polypeptide or protein.

Fourier transform ion cyclotron mass spectrometry, essentially as described in US Patent No. 3,937,955 is also of particular use in the analysis and identification of a

protein isolated using the methods of the present invention. An ion cyclotron uses a fixed magnetic field to deflect an ion of known mass moving at a velocity through the field. Thus, if the magnetic field strength is known, measurements of the ion cyclotron frequency suffices to determine the m/z ratio. In other words, in a static magnetic field the mass-to-change ratio is uniquely determined by the ion-cyclotron frequency. In effect, the static magnetic field converts ionic mass into a frequency analog.

In a preferred embodiment, a mass spectrometer is used to determine the amino acid sequence of a peptide, polypeptide or protein, using, for example, MS/MS. In accordance with this embodiment, an ion of interest is passed into a chamber of a mass spectrometer (a "collision chamber"), where the ion interacts with a gas, such as, for example nitrogen or argon. This interaction with a gas causes fragmentation of an ion, preferably within the peptide backbone. Accordingly, following cleavage of an ion, mass analysis of resulting fragments that differ in mass by a single amino acid from another fragment enables the determination of an amino acid sequence of the ion of interest. Particularly preferred mass spectrometers for the analysis of an amino acid sequence of a peptide, polypeptide or protein are, for example, a triple quadripole (essentially as described in Hunt *et al, Proc. Natl. Acad. Sci. USA, 83*, 6233-6237, 1986), quadripole-TOF (essentially as described in Morris *et al, Rapid Commun. Mass Spectrom., 10*, 889-896, 1996) or MALDI-QqTOF (essentially as described in Loboda *et al, Rapid Commun. Mass Spectrom. 14*, 1047 – 1057, 2000)

The sequence of several overlapping ions may then be assembled, such that the sequence of a region, or even an entire polypeptide or protein is known. Alternatively, the sequence of each individual ion may be used in further analysis.

Following determining the sequence of at least a peptide derived from an immunogenic protein, this sequence is compared to a database of sequences in order to determine whether or not the derived sequence is identical to or homologous to a known sequence. Such a database is available, for example at NCBI and ExPASY or Swiss-Prot. Furthermore, as a mass spectrometer also determines the mass of a peptide, polypeptide or protein, this information is also useful in identifying an immunogenic protein, such as, by comparison to a protein mass library, such as, for example, that provided by the UK Human Genome Mapping Project Resource Centre.

As used herein the term "ExPASY" shall be taken to mean the Expert Protein Analysis System at the Swiss Institute of Bioinformatics at Basel University 4056, Basel, Switzerland.

As used herein the term "Swiss-Prot" shall be taken to mean the protein sequence database of the Swiss Institute of Bioinformatics at Basel University 4056, Basel, Switzerland.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

The present invention is further described with reference to the following non-limiting examples.

Example 1

Sample preparation and immunoglobulin isolation

1.5ml of patient serum stored at −80°C was thawed at room temperature then applied to a 2ml column of protein G–sepharose (Amersham Biosciences, Castle Hill, New South Wales, Australia), previously equilibrated with 20mM phosphate buffer pH7 and incubated on ice for 30 minutes with occasional inversion. The mixture was spun at 6000g for 10 minutes at 4°C and the supernatant decanted. The sepharose pellet was washed with 20mM phosphate buffer. The IgG bound to the sepharose was eluted by addition of 50mM glycine pH2.7 for 20 minutes. After centrifugation as above, the supernatant was discarded and the glycine step repeated. The supernatant was collected from this second glycine elution and stored at −80°C.

Example 2

Isolation of an immunogenic proteins or peptides

A Bradford protein assay is performed on the thawed eluate and thirty milligrams of the immunoglobulin fraction loaded onto a Sephacryl S-200 high resolution gel filtration column (Amersham Biosciences). Fractions ranging from 3000 to 140000 MW are collected, excluding the 150000 IgG fraction. These fractions are pooled and precipitated with 10 volumes of cold acetone at -20°C for 48h then centrifuged at 5000g for 20mins at 4°C. The precipitates are resolubilised in 1-2mls of sample buffer containing 5M urea, 2M thiourea, 2% CHAPS, 2% SB3-10 and 40mM Tris, then simultaneously reduced with 5mM tributyl phosphine and alkylated with 10mM acrylamide for 1h. Samples are aliquoted into 250μl aliquots and stored at −80°C.

Example 3

2 Dimensional gel electrophoresis of immunogenic proteins or peptides and immunoglobulins

The protein content of the samples was estimated using a Bradford assay. Samples were diluted to 2mg/ml with sample buffer as above replacing 40mM Tris with 5mM Tris.

Prior to rehydration of IPG strips, samples were centrifuged at 21000 x g for 10minutes. The supernatant was collected and 10µl of 1% Orange G (Sigma) per ml added as an indicator dye.

First Dimension

Dry 11cm IPG strips (Amersham-Biosciences) were rehydrated for 16-24 hours with 180µl of protein sample. Rehydrated strips were focussed on a Protean IEF Cell (Bio-Rad, Hercules, CA) or Proteome System's IsoElectrIQ electrophoresis equipment for approx 140 kVhr at a maximum of 10 kV. Focussed strips were then equilibrated in urea/SDS/Tris-HCl/bromophenol blue buffer.

Second Dimension

Equilibrated strips were inserted into loading wells of 6-15% (w/v) tris-acetate SDS-PAGE pre-cast 10cm x 15cm GelChips (Proteome Systems, Sydney Australia). Electrophoresis was performed at 50mA per gel for 1.5 hours, or until the tracking dye reached the bottom of the gel. Proteins were stained using SyproRuby (Molecular Probes). Gel images were scanned after destaining using an AlphaImager System (Alpha Innotech Corp.). Gels were then stained with Coomassie G-250 to assist visualisation of protein spots in subsequent analyses.

An example Gel Image is shown in Figure 1.

Example 4

Identification of immunogenic peptides or proteins

A number of proteins were observed in serum samples obtained from subjects suffering from a tuberculosis infection. These proteins were then identified using mass spectrometry.

Prior to mass spectrometry it was necessary to prepare protein samples by in-gel tryptic digestion. Protein gel pieces were excised, destained, digested and desalted using an Xcise™, an excision/liquid handling robot (Proteome Systems, Sydney, Australia and Shimadzu-Biotech, Kyoto, Japan) in association with the Montage In-Gel Digestion Kit (developed by Proteome Systems and distributed by Millipore, Billerica, Ma, 01821, USA). Prior to spot cutting, the 2-D gel was incubated in water to maintain a constant size and prevent drying. Subsequently, the 2-D gel was placed on the Xcise, a digital

image was captured and the spots to be cut were selected. After automated spot excision, gel pieces were subjected to automated liquid handling and in-gel digestion. Briefly, each spot was destained with 100 μl of 50% (v/v) acetonitrile in 50 mM ammonium bicarbonate. The gel pieces were dried by adding 100% acetonitrile, the acetonitrile was removed after 5 seconds and the gels were dried completely by evaporating the residual acetonitrile at 37°C. Proteolytic digestion was performed by rehydrating the dried gel pieces with 30 μl of 20 mM ammonium bicarbonate (pH 7.8) containing 5 μg/mL modified porcine trypsin and incubated at 30°C overnight.

Ten μl of the tryptic peptide mixture was removed to a clean microtitre plate in the event that additional analysis by Liquid Chromatography (LC) - Electrospray Ionisation (ESI) MS was required.

Automated desalting and concentration of tryptic peptides prior to MALDI-TOF MS was performed using C18 ZipTip (Millipore, Bedford, MA). Adsorbed peptides were eluted from the tips onto a 384-position MALDI-TOF sample target plate (Kratos, Manchester, UK or Bruker Daltronics, Germany) using 2 μl of 2 mg/ml α-cyano-4-hydroxycinnamic acid in 90% (v/v) acetonitrile and 0.085% (v/v) TFA.

Digests were analysed using an Axima-CFR MALDI-TOF mass spectrometer (Kratos, Manchester, UK) in positive ion reflectron mode. A nitrogen laser with a wavelength of 337 nm was used to irradiate the sample. The spectra were acquired in automatic mode in the mass range 600 Da to 4000 Da applying a 64-point raster to each sample spot. Only spectra passing certain criteria were saved. All spectra underwent an internal two point calibration using an autodigested trypsin peak mass, m/z 842.51 Da and spiked adenocorticotropic hormone (ACTH) peptide, m/z 2465.117 Da. Software designed by Proteome Systems, as contained in the web-based proteomic data management system BioinformatIQ™ (Proteome Systems), was used to extract isotopic peaks from MS spectra.

Protein identification was performed by matching the monoisotopic masses of the tryptic peptides (i.e. the peptide mass fingerprint) with the theoretical masses from protein databases using IonIQ database search software (Proteome System Limited, North Ryde, Sydney, Australia). Querying was done against the non-redundant SwissProt (Release 40) and TrEMBL (Release 20) databases (June 2002 version), and protein identities were ranked through a modification of the MOWSE scoring system.

Propionamide-cysteine (cys-PAM) or carboxyamidomethyl-cysteine (cys-CAM) and oxidized methionine modifications were taken into account and a mass tolerance of 100 ppm was allowed.

Miscleavage sites were only considered after an initial search without miscleavages had been performed. The following criteria were used to evaluate the search results: the MOWSE score, the number and intensity of peptides matching the candidate protein, the coverage of the candidate protein's sequence by the matching peptides and the gel location.

In addition, or alternatively, proteins were analysed using LC-ESI-MS. Tryptic digest solutions of proteins (10 µl) were analysed by nanoflow LC/MS using an LCQ Deca Ion Trap mass spectrometer (ThermoFinnigan, San Jose, CA) equipped with a Surveyor LC system composed of an autosampler and pump. Peptides were separated using a PepFinder kit (Thermo-Finnigan) coupled to a C18 PicoFrit column (New Objective). Gradient elution from water containing 0.1% (v/v) formic acid (mobile phase A) to 90% (v/v) acetonitrile containing 0.1% (v/v) formic acid (mobile phase B) was performed over a 30-60-minute period. The mass spectrometer was set up to acquire three scan events - one full scan (range from 400 to 2000 amu) followed by two data dependant MS/MS scans.

Proteins were identified using TurboSequest (Thermo-Finnigan) software. Peptides were identified from MS/MS spectra in which more than half of the experimental fragment ions matched theoretical ion values, and gave cross-correlation (a raw correlation score of the top candidate peptide), delta correlation (difference in correlation between the top two candidate peptides) and preliminary score (raw score used to rank candidate peptides) values greater than 2.2, 0.2, and 400, respectively.

Using this method a 49.7kDa protein was identified in the immunoglobulin fraction of a TB subject. This protein was analysed using MALDI-TOF MS and fragments (SEQ ID NO: 2-7) shown in Table 1 identified.

Table 1

Fragments identified from an immunoglobulin associated protein

Peptide Number	SEQ ID NO:	Site of peptide in protein	Sequence
1	2	204-221	FEAVKGECNMGQQEIGFR
2	3	241-255	EIADQHGKSLTFMAK
3	4	305-318	EFTLCYAPTINSYK
4	5	343-351	VVGHGQNIR
5	6	401-419	LPVTLADAAVLFEDSALVR
6	7	436-450	VELAAFNAAVTDWER

Using this information sequence databases were searched and it was predicted that the isolated protein spot was a hypothetical *M. tuberculosis* glutamine synthetase protein (SEQ ID NO: 1). Clearly this demonstrates that this method is of use in the isolation and identification of proteins that are of particular use in the diagnosis/prognosis, treatment or prophylaxis of disorders such as infections or autoimmune diseases.

Example 5

Alternate method of sample preparation and immunoglobulin isolation

1.5ml of patient serum stored at –80°C was thawed at room temperature then applied to a 2ml column of protein G–sepharose (Amersham Biosciences), previously equilibrated with 20mM phosphate buffer pH7 and incubated on ice for 30 minutes with occasional inversion. The mixture was spun at 6000g for 10 minutes at 4°C and the supernatant decanted. The sepharose pellet was washed with 20mM phosphate buffer. The IgG bound to the sepharose was eluted by addition of 50mM glycine pH2.7 for 20 minutes. After centrifugation as above, the supernatant was discarded and the glycine step repeated. The supernatant was collected from this second glycine elution and stored at –80°C.

Following sample preparation, proteins are separated using two-dimensional electrophoresis and immunogenic proteins identified using MALDI-TOF MS essentially as described in Examples 2, 3 and 4.

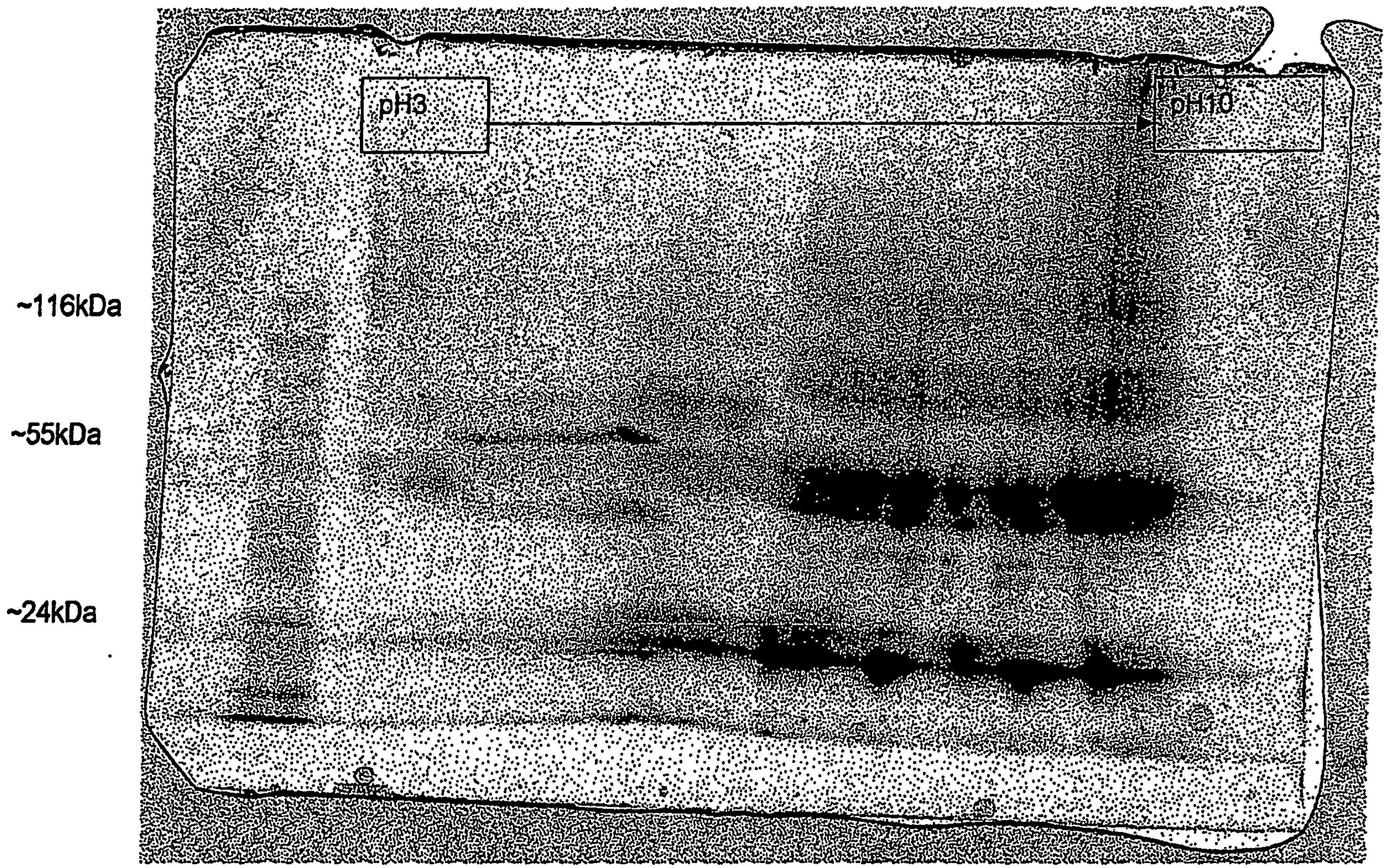

Figure 1